"Express Mail" Mailing Label No.:EV 079 487 595 US

February 6, 2004
Date of Deposit

Our Case No. 6270/134

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
APPLICATION FOR UNITED STATES LETTERS PATENT

INVENTORS: John Christopher Van Gorp
J. Bradford Forth
Theresa M. Köster
Mark P. G. Kowal
Daniel W. O'Neill
Daniel J. Wall

TITLE: A METHOD AND SYSTEM FOR CALCULATING AND DISTRIBUTING UTILITY COSTS

ATTORNEY: James L. Katz (Reg. No. 42,711)
BRINKS HOFER GILSON & LIONE
POST OFFICE BOX 10395
CHICAGO, ILLINOIS 60610
(312) 321-4200

A METHOD AND SYSTEM FOR CALCULATING AND DISTRIBUTING UTILITY COSTS

RELATED APPLICATIONS

[0001] This application claims the benefit of the filing date under 35 U.S.C. § 119(e) of U.S. Provisional Application Serial No. 60/455,881, entitled ENERGY ANALYTICS FOR AN ENERGY DISTRIBUTION SYSTEM, filed February 7, 2003, which is hereby incorporated by reference, and U.S. Provisional Application Serial No. 60/455,788, entitled HUMAN-MACHINE INTERFACE FOR AN ENERGY ANALYTICS SYSTEM, filed February 7, 2003, which is hereby incorporated by reference.

SUMMARY

[0002] The present invention is defined by the following claims, and nothing in this section should be taken as a limitation on those claims. By way of introduction, the preferred embodiments described below relate to a rate engine for use in a utility distribution system. The rate engine includes: an input module operative to accept utility data, rate data and time data, the time data including at least one logging interval; a processing module coupled with the input module and operative to compute at least one cost based on the utility data and rate data, the at least one cost being associated with the at least one logging interval; and an output module coupled with the processing module and operative to output the at least one cost.

[0003] The preferred embodiments further relate to a method of calculating the per logging interval cost of a utility. In one embodiment, the method includes: receiving utility data, rate data and time data, the time data including at least one logging interval; computing at least one cost based on the utility data and rate data, the at least one cost being associated with the at least one logging interval; and outputting the at least one cost.

[0004] Further aspects and advantages of the invention are discussed below in conjunction with the preferred embodiments.

BRIEF DESCRIPTION OF THE DRAWINGS

[0005] Figures 1a and 1b each depict an overview of a system having a rate engine according to different embodiments;

[0006] Figure 2 depicts a time scale for demonstrating time intervals of relevance to the disclosed embodiments;

[0007] Figure 3 depicts a first exemplary method for distributing costs;

[0008] Figures 4a and 4b depict a second exemplary method for distributing costs;

[0009] Figures 5a and 5b depict a third exemplary method for distributing costs;

[0010] Figure 5c depicts a fourth exemplary method for distributing costs;

[0011] Figure 5d depicts a fifth exemplary method for distributing costs;

[0012] Figures 6a, 6b, 6c and 6d depict exemplary tables outputted by the disclosed embodiments; and

[0013] Figures 7a, 7b, 7c and 7d depict sample cost sensitivity scenarios.

DETAILED DESCRIPTION OF THE PRESENTLY PREFERRED EMBODIMENTS

[0014] Billing engines typically return a cost for a billing period (usually a month). This is useful for shadow-billing applications. However, this is not useful for real-time bill tracking applications, ad-hoc queries, calculating the cost of production runs, real-time pricing and other such applications.

[0015] The disclosed embodiments integrate aspects of the Instantaneous Cost Engine, Inference Engine, Distributed Enterprise Rate Engine and Incremental Rate Engine described in U.S. Provisional Application Serial No. 60/455,881, captioned above. The disclosed embodiments comprise a Rate Engine, which may execute on a computer or other processor, which provides cost analysis with a finer granularity than monthly or billed costs. The Rate Engine receives inputs such as utility data and rate data, and generates cost data on a per logging interval basis. A logging interval is the finest sampling frequency available, which varies depending on the application, such as from microseconds to minutes or hours. The resulting cost breakdown allows the user understand and analyze their energy costs in the same manner that they analyze other costs of business such as labor and goods. For example, the interval costs can be aggregated to obtain a total bill cost, or a cost associated with time periods other than a

billing interval. This is useful for calculating the cost of a shift or production run. The interval costs may also be used as inputs for further analysis.

Inputs to the Rate Engine

[0016] FIG. 1a shows an exemplary system having a rate engine according to one embodiment. In particular, FIG. 1a shows a Utility 100, as defined below, being used / consumed, and a Measuring Device 102 coupled with the Utility 100 so as to measure the usage/consumption of the Utility 100. The Measuring Device 102 is further coupled with Network 104, which is further coupled with Communication Interface 105 of Rate Engine 106, and Data Storage 108. Data, such as usage/consumption measurements, collected by the Measuring Device 102 is transported over the Network 104 to the Rate Engine 106, or to some other data storage location 108 from which it may be retrieved by the Rate Engine 106 at a later time. Additionally, one or more Utility Management Entities 110 that manage the provisioning of Utility 100 are coupled with the Utility 100. The Utility Management Entities 110 may publish rate data via a Rate Data Server 112, which is coupled with the Network 104. The Communication Interface 105 may receive rate data from the Rate Data Server 112 or Data Storage 108 or a database or web service (not shown) or some other publication service (not shown) via the Network 104.

[0017] Herein, the phrase "coupled with" is defined to mean directly connected to or indirectly connected through one or more intermediate components. Such intermediate components may include both hardware and software based components.

[0018] The Utility 100 may be any product or service used or consumed over a period of time, and where actual usage/consumption at any given instant may vary, and further wherein consumption/usage is measured and billed according to overall consumption/usage and/or consumption/usage at a given instant. Exemplary Utilities 100 may include Water, measured in liters for example, Air, measured in liters for example, Gas, measured in liters for example, Electricity, measured in kilowatts ("KWs") for example, Steam, measured in liters for example, Emissions, measured in liters or grams for example, Bandwidth, measured in bits per second for example, computer processing power, measured in Million Instructions Per Second ("MIPS") for example, and so forth. It will be appreciated that the actual units of measurement may vary depending upon the

Utility 100 being measured and the infrastructure implementation and cost structure of delivery of that Utility 100. For example, the Rate Engine 106 could be used in an energy management system to distribute the costs associated with energy usage across time intervals. The Measuring Device 102 may be any device that is coupled with the Utility 100, and is able to measure usage/consumption of the Utility 100 on a periodic basis, such as an Intelligent Electronic Device ("IED"). The network 104 may include public or third-party operated networks such as: Virtual Private Networks ("VPNs"), Local Area Networks ("LANs"), Wide Area Networks ("WANs"), telephone, dedicated phone lines (such as ISDN or DSL), Internet, Ethernet, paging networks, leased line; Wireless including radio, light-based or sound-based; Power Line Carrier schemes and so forth. Communication Interface 105 can be any suitable software or hardware required by the Rate Engine 106 to receive data from the Network 104. For example, where the Rate Engine 106 is a self-contained device, the Communication Interface 105 incorporates all hardware and software needed to connect the device to the Network 104. In one embodiment, the Rate Engine 106 may be implemented as a hardware device, in an alternate embodiment the Rate Engine 106 is a software product with program modules that run on a personal or mainframe computer with supporting hardware/software and network connections. Data Storage 108 may include a database, a data acquisition module or some other data server. The Utility Management Entities 110 may include utility companies such as electric, water or gas suppliers, energy marketers, landlords, property managers, cost center managers, energy managers, key account representatives, and so forth. The Rate Data Server 112 includes any source of rate data, as defined herein, including a database, a website, an Excel spreadsheet used with the Microsoft Excel spreadsheet program, published by Microsoft Corporation, located in Redmond, Washington, a handwritten document, and so forth.

[0019] IEDs include revenue electric watt-hour meters, protection relays, programmable logic controllers, remote terminal units ("RTUs"), fault recorders, other devices used to monitor and/or control electrical power distribution and consumption, RTUs that measure water data, RTUs that measure air data, RTUs that measure gas data, and RTUs that measure steam data. IEDs are widely available that make use of memory and microprocessors to provide increased versatility and additional functionality. Such

functionality includes the ability to communicate with other hosts and remote computing systems through some form of communication channel. IEDs also include legacy mechanical or electromechanical devices that have been retrofitted with appropriate hardware and/or software allowing integration with the power management system. Typically an IED is associated with a particular load or set of loads that are drawing electrical power from the power distribution system. The IED may also be capable of receiving data from or controlling its associated load. Depending on the type of IED and the type of load it may be associated with, the IED implements a function that is able to respond to a command and/or generate data. Functions include measuring power consumption, controlling power distribution such as a relay function, monitoring power quality, measuring power parameters such as phasor components, voltage or current, controlling power generation facilities, computing revenue, controlling electrical power flow and load shedding, or combinations thereof. For functions that produce data or other results, the IED can push the data onto the network to another IED or back end server/database, automatically or event driven, or the IED can send data in response to an unsolicited request. IEDs capable of running Internet protocols may be known as "web meters". For example, a web meter may contain a web server.

[0020] FIG. 1b shows an exemplary system for use with a rate engine according to another embodiment. The Rate Engine 106 receives various inputs 120, including, Utility Data 122, Rate Data 124, Time Data 126 and optional Meta Data 128, described in more detail below. The inputs 120 may be retrieved from various sources such as hardware devices, databases, web services and so forth, such as the Measuring Device 102, Data Storage 108 or Rate Data Server 112 of FIG. 1a. The inputs 120 are received by the Input Module 132 of the Rate Engine 106. The Input Module 132 is coupled with the Processing Module 134, and passes inputs to the Processing Module 134. The Processing Module 134 combines the Utility Data 122 and Rate Data 124 to create costs associated with the Time Data 126, as will be described in greater detail below. The Processing Module 134 is coupled with the Output Module 136. The Processing Module 134 transmits the results of the calculated and distributed costs to the Output Module 136, which is responsible for providing third party applications with the calculated and distributed costs. This can be done by transmitting the output 140 to a particular

destination, by making it available for access such as on a web server, or via an application programming interface ("API"). Some examples of the calculated and distributed costs are show by output 140. The output 140 may be one or more time intervals, with a cost associated with each time interval. Two exemplary outputs 142 and 144 are shown. The outputs 140 may then be used to perform further analysis 150. These outputs 140 will be described in greater detail below.

[0021] Utility Data 122 may be broken down into resources. For example, where the Utility Data 122 is for electricity, the resources may be usage/ KWH, demand / KW, power factor, KVAR, KVH and so forth. Where the Utility Data 122 is for emissions, the resources may be credits and debits. The Utility Data 122 may be real data, or may be hypothetical data. Hypothetical data may represent data in the past, data in the future, data that has been scaled, data that has been shifted, data that has been edited, data that has been estimated/interpolated, data that has been normalized and/or data that has been modeled. Hypothetical data can be used to run what-if scenarios, generate forecasts and to correct missing or false data. What-if scenarios allow a user to quickly answer questions such as "If a particular motor was turned on at a different time, how would this affect my energy bill?" Scaled data may represent a scenario where either more or less of the resource is used, for example, reducing lighting by 5%. Shifted data may represent a scenario where resource usage is shifted to a different time interval. Editing the data may represent a scenario where an error in the original measurement is corrected. Estimating or interpolating the data may represent a scenario where the original data is not available, and must be estimated or interpolated based on previous usage patterns or some other model. The data can be normalized with respect to such dimensions as temperature, square footage, occupancy, etc. A dimension is a way of classifying information. The data can be modeled using multivariate regression, neural networks, Fast Fourier Transforms ("FFT") and so forth.

[0022] Rate Data 124 defines how resource usage will be billed. Resource usage billing may be defined in a tariff, or may be based on a real-time pricing scheme. A tariff is a contract or document that defines how a customer will be billed for the resources they have used. Tariffs may be retrieved from a utility, database, web service, or some other source. Real-time pricing schemes are pricing models where the cost of a resource

changes frequently and is based on supply and demand, for example, one example is the Independent Market Operator's hourly Ontario, Canada energy price. Real-time prices may be retrieved from a web service or some other source. When the Rate Engine 106 is calculating costs for a large number of time intervals, it may require that the Rate Data 124 include multiple tariffs, or combinations of real-time pricing values and tariffs.

[0023] Tariffs may be composed of a number of different charges. Charges are different from costs, which are what the customer pays. Which charges appear on a bill will vary based on the agreement reached between the utility and customer. Some typical charges appearing on an electricity tariff include Time Of Use ("TOU"), demand, usage, power factor and customer charge. Other charges include transmission and distribution, fuel, and access, government, state or municipal fees or taxes, service charges, credits (payback for previous overcharging) and debt retirement charges (paying off publicly owned utility debt.)

[0024] Examples of demand charges include contract demand, demand peak with ratchet and billing demand. Contract demand is where the customer has a contract for a demand of X. Where the customer's power use is less than the contract, the customer must still pay some percentage of the contract, usually between 50% to 100%. Demand Peak with Ratchet is where each new peak affects future billing periods for some time frame, which can be a year or more. That is, in future bills, even where actual demand of that billing period is less than the previous peak, some percentage of that previous peak has to be paid. Billing demand can be actual demand, contract demand, or percentage of peak from previous Y billing periods.

[0025] Examples of penalty charges include demand or KW, lagging power factor surcharge and penalties for failing to comply with request to interrupt usage. A lagging power factor surcharge is billed when the power factor goes below certain percentage for period of time.

[0026] Some charges may have a Time-Of-Use ("TOU") element as well, which dictates varying costs for the charges, based on the time and day the charge is incurred. For example, a tariff may indicate that between 6 am to 6 pm on Monday to Friday that energy costs 10¢ per KWH, and that the rest of the time energy costs 7¢ per KWH.

[0027] Some tariffs include a price tier charge. Price tiers may also be known as penalty points, price steps or ratchets. A price tier stipulates that over a certain usage or demand the charge per unit will be higher. For example, a tariff may indicate that where demand < 12 KW, the price per KW is $2, and where demand >= 12 KW, the price per KW is $5. In another example, the tariff may indicate that where the usage per billing period is < 15000 KWH, the cost per KWH is 5¢ and that once usage for the billing period exceeds 15000 KWH, each additional KWH will cost 7¢ each.

[0028] Some charges such as demand charges, TOU, taxes, customer charge are usually billed per month / billing period, whereas real-time prices are usually published more frequently, for example, once an hour or once a minute. The Rate Engine 106 must normalize these different costing units so that they can be compared, aggregated and disaggregated.

[0029] Future events may affect previously calculated interval charges. For example, forecasted real time prices that decisions are based on can change after the fact, or the setting of a new demand peak can affect previous weeks. Each per interval cost may need to be recalculated after each logging interval.

[0030] The Time Data 126 defines a time interval composed of one or more logging intervals, where a discrete cost is to be assigned to each of the logging intervals within the time interval. The time interval may describe any time period. Time intervals will be described in greater detail below.

[0031] The optional Meta Data 128 includes any other data that is used to describe, categorize or otherwise identify the Utility Data 122, Rate Data 124 and Time Data 126. Meta Data 128 can include billing period id, cost center id, geographic location, address or ZIP code. Meta Data 128 may be output along with Outputs 140.

[0032] The Rate Engine 106 distributes charges defined by the Utility Data 122 and Rate Data 124 to the logging intervals defined by the Time Data 126. The section below describes how this distribution is performed.

[0033] Where the Time Data 126 defines only one logging interval, the Output 140 of the Rate Engine 106 includes the time interval and the cost associated with that time interval, as is shown in table 142. Where the Time Data 126 defines more than one logging interval, the Output 140 includes multiple logging intervals and their associated

costs as is shown in table 144. It will be appreciated that additional data, such as Meta Data 128 may form part of the output 142, 144, but is not shown here for simplicity. As will be described in greater detail below, the Outputs 140 of the Rate Engine 106 are not limited to the two dimensional tables 142 and 144 depicted in FIG. 1, but can be multi-dimensional. By varying the quantity of resource consumed during a single logging interval, or by calculating costs for multiple resources, the Outputs 140 of the Rate Engine 106 can become multi-dimensional.

[0034] The Outputs 140 of the Rate Engine 106 are available to be used for some further Analysis 150. This analysis may consist of what-if scenarios, cost allocation, sensitivity analysis and so forth, which will be described in greater detail below.

Time Intervals

[0035] Referring now to FIG. 2, an exemplary time scale 160 is depicted for demonstrating time intervals of relevance to the disclosed embodiments. Time scale 160 has times t0 to t10 marked on it. Time t0 marks the start of the year. Time t1 marks the start of billing period 1. Time t5 marks the start of billing period 2. Time t7 marks the start of billing period N. Time t8 is the current time. Time t9 marks the end of billing period N, and time t10 marks a time years later from time t0.

[0036] Times t2 and t3 define one logging interval, which is the smallest granularity of time defined by the time scale 160. The logging interval defined by times t2 and t3 may affect future billing periods.

[0037] Times t1 to t5 define one billing period (billing period 1). Note that billing periods do not necessarily line up with monthly boundaries.

[0038] Times t3 to t4 define a subset of billing period 1, and times t4 to t6 define a time interval spanning parts of two billing periods (billing period 1 and billing period 2). Such time intervals may be of interest because they represent a shift, a production run, a process change, a government regulation change, a tenant change, a daylight savings transition, a season change, and so forth. The disclosed embodiments allow costs for production runs or other loads to be accurately calculated even when they don't coincide with a billing period boundary.

[0039] Times t7 to t8 define a time interval that is equivalent to the first portion of a current billing period. Such a time interval may be useful when requesting the current bill to date, answering the question "how much have I spent so far on this billing period?" This allows the user to adjust their resource usage on the fly in response to current conditions.

[0040] The concept of time intervals described herein is important in understanding how the Rate Engine 106 functions. As discussed previously in relation to FIG. 1b, the Rate Engine 106 takes various Inputs 120, and distributes the charges defined by the Utility Data 122 and Rate Data 124 to the logging intervals defined by the Time Data 126. However, it is likely that the logging intervals defined by the Time Data 126 are affected by logging intervals which are outside of Time Data 126. Therefore, the Rate Engine 106 must intelligently decide what the spanning interval is, based on the applicable rules. A spanning interval contains all time intervals that affect or are affected by the logging intervals of interest in a particular run of the Rate Engine 106. For example, the rate data may reference, be based on or affect previous or future billing periods. In this case the affected billing periods are retrieved and used in calculating the distribution.

How the Rate Engine Distributes Costs

[0041] To provide costing data on a per interval basis, the charges must first be distributed across the intervals. There are many ways to do this and the type of approach used to distribute will depend on the type of charge being distributed. Some charges may be easily divided among intervals, whereas for others it may not be obvious how their contribution to the bill can be distributed. It will be appreciated that different charges may be distributed using different algorithms, depending on the situation and user requirements.

[0042] In one example, charges are distributed to the time interval they occurred in. For example, if a penalty event occurs at a given time interval, the charge for that event will be assigned only to that time interval, and will not be shared with any other. In another example, where a penalty event occurs at a given time interval and has repercussions for future time intervals, the future time intervals will take the per interval

cost. In another example where there is a usage price tier, earlier intervals will have the cheaper rate and later intervals will have the more expensive rate.

Flat Distribution

[0043] In an alternate example, a charge for a billing interval is distributed evenly across all of the intervals that the bill applies to. For example, a demand charge can be divided by the number of logging intervals, and the result of this division is distributed evenly among the logging intervals. This method is known as Flat Distribution, and may be more suitable for certain "flat" charges, such as customer charge.

[0044] It will be appreciated that there may be other approaches for computing a Flat Distribution, and that the foregoing is presented merely by way of example.

Weighted Distribution

[0045] In an alternate example, one or more charges are distributed based on weighting of another charge. For example, charges such as penalties, taxes, customer charges and so forth can be distributed based on interval usage, weather, apparent power, reactive power, real power and so forth. This technique for distributing charges is called Weighted Distribution. Referring now to FIG. 3, an example of the Weighted Distribution algorithm is shown. Graph 205 depicts five time intervals on the X-axis (t1, t2, t3, t4 and t5) and penalty charges are depicted on the Y-axis. In this example, a penalty event (say for KW) occurred in time interval t3, and a related charge of $100 has been billed to t3. Graph 210 shows the usage for the five time intervals (say usage is KWH). Depending on the tariff(s) governing t1 through t5, it may be the case that if t3 had not caused a penalty event, that t4 would have, or t1 or t5, or t2. Under a weighted distribution, the penalty charge of $100 is distributed to the five time intervals based on usage. The resulting distribution is shown in graph 215, where each time interval is given a charge that is proportional to the usage of that interval with respect to the other four intervals. This is useful, because it is possible that had t3 not incurred the penalty, that t4 would have.

[0046] It will be appreciated that there may be other approaches for computing a Weighted Distribution, and that the foregoing is presented merely by way of example.

Zeroing Distribution

[0047] In an alternate example, where a total cost for a bill or some other time interval is known, the cost of an interval is determined by arranging the resources such that if calculated on their own, the cost for that interval would be $0. Then, a new total cost of the bill is recalculated, and the assigned cost of the interval is the difference between the old and new total costs. It will be appreciated that either all resources can be manipulated together, or they can be manipulated separately, for different results, noting also that some resources are ratiometrically linked, such that changing one resource will have an affect on another resource. This technique for distributing charges is called Zeroing Distribution.

[0048] Referring now to FIG. 4a and 4b, an example of the Zeroing algorithm is shown. Table 405 shows the complete resource information (resource 1 to resource N) for time intervals t1 to tM. The total cost for t1 through tM is $P. Table 410 shows that the resources for t1 have been arranged such that if t1 was calculated on its own, the total cost would be $0. This doesn't mean that the value of the resource is always going to be zero however. For example, if power factor is a resource, then the value for the power factor resource would be set to 1, where current and voltage are in sync, rather then to 0, where they are opposed and no useful power is being transmitted. Now that t1 has been zeroed, the total cost for t1 to tM is calculated, yielding a total cost of $A. The marginal cost of t1 is thus the difference between $P and $A. Table 415 shows the same process for t2, where the t2 resources are zeroed, a total cost $B is calculated, and a marginal cost ($P - $B) for t2 is determined. This same process is repeated for time intervals t3 to tM. Table 420 shows the resulting cost associated with each time interval. As is indicated by item 425, in some cases when the costs of the discrete intervals are added up, the result is not the same as the original bill. In this case a scaling factor can be applied to each interval cost, as is indicated in Table 430 so that the costs will add up to the billed total.

[0049] It will be appreciated that there may be other approaches for computing a Zeroing Distribution, and that the foregoing is presented merely by way of example.

Slicing Distribution

[0050] In an alternate example, a charge that has been applied to one interval is shared between other intervals that might have incurred the charge if the situation had been slightly different. For example, a peak demand charge can be distributed among all the intervals that had a higher than average demand. In this approach, the charge is divided into subparts, and the subparts are distributed evenly between all intervals contributing to that subpart. This approach for distributing charges is called Slicing Distribution.

[0051] This approach will be better understood by referring to FIG. 5a and FIG. 5b. Chart 505 depicts a sample resource usage for time intervals t1 to t10. The first step in the technique is shown in chart 510, where the intervals are sorted in order of usage. The next step, shown in chart 515, is to take a slice across all intervals which is the same and no greater than the interval with the lowest usage, such that the usage of each interval is set to the usage of the interval with the lowest usage. The cost of the total interval is now calculated given these values. The resulting cost is distributed evenly among all intervals. In chart 515, the total cost for the slice is $100, there are 10 intervals, so each interval is distributed $10. The first slice removed one or more intervals that had the lowest usage. In the next step, shown in chart 520, a slice is taken across all remaining intervals which is the same and no greater than the interval with the lowest usage among the remaining intervals. The usage of the remaining intervals is set to the usage of the remaining interval with the lowest usage and the intervals not in the slice keep their actual value. Now the cost of all intervals is calculated given these values. The cost of the previous slice(s) is/are subtracted from the new cost, and the remainder is evenly distributed to all intervals involved in this slice. For example in chart 520, the total cost is $260, subtracting the $100 of the previous slice leaves $160 to be distributed among 8 intervals, giving $20 to each interval. The same process is repeated until the total usage of the total interval has been calculated, an example of which is shown in charts 525, 530, 535, 540, 545. Now the incremental costs of each interval are added up to result in a cost distributed to each interval. The distributed interval costs for each interval are shown in chart 550. Interval t6 only participated in the first slice, so is distributed $10. Interval t10 participated in the first and second slices, so is distributed $30 ($10 + $20). In the case where the sum of the distributed interval costs does not sum to the billed value, some sort

of scaling can be performed on the interval costs to make them add up to the billed value. This strategy makes the per interval cost related to a penalty apparent. For example, where the penalty is a demand peak, it allows the user to see what the associated cost saving would be of shaving off this peak.

[0052] Alternately, Slicing Distribution can be carried out using a top down approach. Using the example in FIG. 5a and FIG. 5b, t4 is sliced first, setting the usage for t4 the same as for t3. The cost for all intervals is now calculated, and the difference between the actual cost and the calculated cost is applied to t4. This process is continued in a top down manner.

[0053] It will be appreciated that there may be other approaches for computing a Slicing Distribution, and that the foregoing is presented merely by way of example.

Slicing Distribution with Price Tiers

[0054] In an alternate example, where there are known price tiers, charges are distributed in a manner similar to the Slicing Distribution, except that an additional slice is made for each price tier. An example of this distribution is shown in FIG. 5c. The tariff covering the time interval t1to t3 of chart 560 indicates that where KW is less than or equal to 12, the cost is $2 per KW, but that when KW is greater than 12, the cost jumps to $5 per KW. KW for t1 is 20, for t2 is 15 and for t3 is 10. The penalty charge for KW is $100 (20KW multiplied by $5). Instead of just making slices at 10 KW, 15 KW and 20 KW to correspond with intervals t1, t2 and t3, an additional slice is made at 12 KW to correspond with the price tier. Otherwise the algorithm is the same as for the basic Slicing Distribution. In the case where the sum of the distributed interval costs does not sum to the billed value, some sort of scaling can be applied to make the interval costs add up to the billed value.

[0055] It will be appreciated that there may be other approaches for computing a Slicing Distribution with Price Tiers, and that the foregoing is presented merely by way of example.

Tiered Distribution

[0056] In an alternate example, where there are known price tiers, the resource charges for each tier can be separately distributed. An example is shown in FIG. 5d. The tariff covering the time interval t1to t3 of chart 570 indicates that where KW is less than or equal to 12, the cost is $2 per KW, but that when KW is greater than 12, the cost jumps to $5 per KW. KW for t1 is 20, for t2 is 15 and for t3 is 10. The penalty charge for KW is $100 (20KW multiplied by $5). If a Weighted Distribution was used to distribute the charges to the three intervals, then t3 would be unfairly charged, as t3 never exceeded the penalty point of 12 KW. The technique is implemented as follows. The cost of the first price tier is calculated. In this case it is $24 (12KW multiplied by $2). The $24 is distributed among all intervals that participated in that price tier using a weighted distribution. In the example of FIG. 5d, in the first tier, t1 and t2 have a usage of 12 units and t3 has a usage of 10 units. (12 + 12 + 10 = 34), so the $24 is divided by 34, resulting in ~$.71 per unit of usage in the first tier. Now for t1, multiply $.71 by 12 units, resulting in $8.47, and continue for each interval. Next, the cost of the following price tier is calculated. In this case it is $76 ($100 minus the $24 allocated to the first tier). The $76 is distributed among all intervals that participated in that price tier, using the same weighted distribution. In this example, only t1 and t2 participated in the second tier. Once all price tiers have been calculated, the cost for each interval is calculated by adding up the component price for each tier for which the interval was a participant.

[0057] It will be appreciated that there may be other approaches for computing a Tiered Distribution, and that the foregoing is presented merely by way of example. It will also be appreciated that in general other ways of distributing costs to the time intervals exist and that the aforementioned approaches were provided merely by way of example.

The Outputs

[0058] Once the charges have been allocated to the intervals, the Rate Engine 106 outputs the tables of time intervals and their associated costs. These costs may be stored in a database, forwarded to another application, such as Excel, manufactured by Microsoft Corporation, located in Redmond, Washington, Crystal Reports, manufactured

by Business Objects of San Jose, California, Cognos, manufactured by Cognos Corporation of Ottawa, Canada, etc, or displayed. The charges of the bill may be displayed in a pie chart format. For example, an electricity bill might show portions of the pie such as fixed charges, peak demand charge, TOU and so forth.

[0059] The costs output by the Rate Engine 106 can also be used by various analysis modules. Such analysis modules might include modeling, normalization, comparison, what-if analysis, forecasting, aggregation, control, cost allocation, best path or sensitivity modules. The visualization, normalization, aggregation, comparison, what-if analysis, sensitivity, control and cost allocation analysis can all be done with respect to a dimension, such as floor space, occupancy, process or production run, shift, sub meter, feeder, time, necessity, bill period, day of week and so forth. The analysis may entail running the Rate Engine 106 again with additional inputs.

[0060] An example of modeling would be to model the cost output of the Rate Engine 106 with respect to an external factor, such as weather, heating degree days, wind, cloud cover, or sunrise/sunset times using some technique such as linear regression.

[0061] An example of normalizing is to normalize costing output with respect to some dimension such as total square footage, leased and occupied square footage, leased and unoccupied square footage, tons of pulp, etc.

[0062] Costing output can be compared with other costs or with other units of data, such as comparing costs to weather or another external factor described above. Hypothetical and real costs can be compared to each other. For example, the Rate Engine 106 can be run with the same utility data but different tariffs, or the same tariff and hypothetical utility data, or different tariffs and different data, and the outputs can be compared. This is useful when the user wishes to perform a what-if analysis, such as: what-if the weather gets hotter, or some plants are closed, or the load profile is changed, or the lighting load is reduced by 10% or the tariff changes etc. It is also useful when making budgeted to actual comparisons. The what-if analysis may be used to forecast costs. For example, a forecasting module can have a forecast containing typical values for each logging interval for the billing interval, a year, or some other time period. Then as the intervals complete, the forecasted intervals can be replaced by the real values, and the yearly bill can be calculated on the fly (combining the real and predicted values), and

alarm bells can be set off if the current interval will significantly affect the total cost. Calculations may be done either after the interval is complete or mid-interval (predicting how the interval will complete based on current and predicted usage, and current trajectory).

[0063] A control action can be made automatically by some intelligent software, or by an operator. Actions include curtailing use, shifting loads and negotiating / procuring excess capacity.

[0064] Cost allocation is the process of creating sub-bills for tenants, cost centers or other financial entities from a total bill. Where the quantity being charged for is electricity, there are various resources, such as energy, demand and power factor that must be allocated to the financial entities. This is done by allocating portions of the bill based on some quantity such as sub metering, square footage, metering and square footage together (this is where a meter is shared by two financial entities), lighting, Heating Ventilating and Air Conditioning ("HVAC"), and so forth.

[0065] Best Path analysis entails given certain constraints, using artificial intelligence to search through the interval data to find the optimal usage level for each interval. However, making a change in one interval may affect the outcomes for other intervals. Calculating the best path involves calculating how usage in one interval will affect what is optimal in other intervals.

Sensitivity Analysis

[0066] In many situations, the user/consumer of resources wants to optimize the price paid for those resources. To do this, they must understand risks and sensitivity. Sensitivity analysis leads to an understanding of how the system responds to perturbations from the steady state. It can also be described as understanding how cost output responds when variations are made, understanding opportunities, and gleaning efficiency gains. Sensitivity can be used as a form of historical pattern recognition to identify times when more or less resource can be used, for future prediction based on past usage, or to help understand the volatility of a business, for example when using a hedging strategy for purchasing.

[0067] The general idea of cost sensitivity is for each interval, vary the resource usage by some increment from some value less than the actual usage to a value greater than actual, and calculate the cost at each increment. Each logging interval may be varied independently or the sensitivity module can vary all, or a subset of, intervals by some value at the same time. The slope for each logging interval is then evaluated within a certain tolerance of actual either mathematically or by graphing the costs in a 3D contour or mesh graph. Depending on the direction of and steepness of the slope, opportunities to reduce usage and save money, increase usage and not pay more for it, or do nothing will become apparent. The best opportunities can then be exploited. This process will now be explained in greater detail.

[0068] The first step is to create the raw data, which includes of a table of time intervals, resource usage increments and associated costs. Where the utility is electricity, examples of resources to do sensitivity analysis with respect to are KWH, KW, and power factor, whereas fixed charges are not sensitive as they don't change. Referring now to FIG. 6a, two sample tables 605 and 610 are shown. In table 605, resource usage for time interval t1 is varied from 0% through 125%, and data points are calculated at 5% increments. In this case the actual cost is at 100%. In table 610, resource usage for time interval t1 is varied from 0 units through 48 units, and data points are calculated at every 2 units. In this case the actual cost is at 38 units. Referring now to FIG. 6b, a larger output table 615 is shown. This table shows a number of time intervals, t1 through t13, with associated cost values for 0% through 125% resource usage. The cost values in the 100% column are the actual interval costs. It will be appreciated that tables 605, 610 and 615 are shown for example purposes only, and that the resource usage can vary between some other lower and upper bound, such as 90% and 110%, or whatever is of relevance to the current user and application. Furthermore, the increment will be determined by the particular situation. The calculations can be done for one or more time intervals. The resource usage units being varied are shown here as percentages or units. The benefit of varying by unit is that similar things are being compared, whereas percentage of total can vary greatly. The outputted range of values can represent cost of the total bill, $/resource, $/resource/interval or some other cost metric. Furthermore, the resource usage that is varied may either be done on a per resource basis or on all resources at once, or on one

resource at a time, for each resource. The output of the last option will be a multi-dimensional graph.

[0069] Tables 605, 610 and 615 all depict discrete dollar costs for each value. Referring now to FIG. 6c and FIG. 6d, an alternate format for presenting the values is shown. The dollar values are broken down into price bands, shown in key 625, which are then color coded or shading coded. Tables 630, 635 and 640 show various sample outputs, where the values are made easier to understand by being shading coded.

[0070] Once the raw data has been generated, the next step is to evaluate cost sensitivity, and identify opportunities. Referring now to FIG. 7a, chart 705 depicts two scenarios for a time interval, considering the situation where usage may be reduced. The two scenarios are shown on the same chart to facilitate their comparison. Chart 705 shows $/interval on the Y-axis and resource usage on the X-axis (this can be for a single resource, or for multiple resources), and also a user defined lower bound Window, indicating that the user under no circumstances wishes to reduce their resource usage below this Window. The Window can be described as a percentage or as a number of units from the actual usage, and acts as a filter. In scenario A the $/interval does not decline appreciably between the actual usage and the lower bound window, whereas in Scenario B there is a steep negative slope between the actual usage and the lower bound window. For Scenario A, the cost sensitivity graph indicates that resource usage shouldn't be reduced, as there would only be minor cost savings. For Scenario B, the cost sensitivity graph indicates that resource usage may be reduced, as significant cost savings would be achieved within the window determined by the user. The cost curve of Scenario B is such that resource usage shouldn't necessarily be reduced all the way to the lower bound window, as the benefits flatten out.

[0071] Referring now to FIG. 7b, chart 710 depicts two scenarios for a time interval, considering the situation where usage may be increased. Chart 710 shows $/interval on the Y-axis and resource usage on the X-axis (this can be for a single resource, or for multiple resources), and also a user defined upper bound Window, indicating that the user under no circumstances wishes to increase their resource usage above this Window. The Window can be described as a percentage or as a number of units from the actual usage, and acts as a filter. In scenario A there is a steep positive slope between the actual usage

and the upper bound window, whereas in Scenario B the $/interval does not increase appreciably between the actual usage and the upper bound window. For Scenario A, the cost sensitivity graph indicates that resource usage shouldn't be increased as the cost would be prohibitive. For Scenario B, the cost sensitivity graph indicates that resource usage may be increased as there is a low cost associated with the increased usage. The cost curve of Scenario B is such that resource usage shouldn't necessarily be increased all the way to the upper bound window, as the cost starts to increase.

[0072] Referring now to FIG. 7c and FIG. 7d, charts 715 and 720 show a derivative of cost. Chart 715 depicts two scenarios for a time interval, considering the situation where usage may be reduced. Chart 705 shows $/resource on the Y-axis and resource usage on the X-axis (this can be for a single resource, or for multiple resources), a user defined lower bound Window, indicating that the user under no circumstances wishes to reduce their resource usage below this Window, and a Minimum Savings Line, indicating that $/resource must be below a certain value to be worth reducing usage. The Window can be described as a percentage or as a number of units from the actual usage. The Window and the Minimum Savings Line both act as filters. In scenario A the $/resource increases as resource usage is reduced. This can occur in a situation where there are price tiers. The $/resource does eventually drop below the Minimum Savings Line, but well outside of the lower bound window. In Scenario B the $/resource does drop below the Minimum Savings Line, but eventually rises up. For Scenario A, the cost sensitivity graph indicates that resource usage shouldn't be reduced, as the $/resource increases. For Scenario B, the cost sensitivity graph indicates that resource usage may be reduced, as cost savings will be achieved within the window determined by the user.

[0073] Referring now to FIG. 7d, chart 720 depicts two scenarios for a time interval, considering the situation where usage may be increased. Chart 720 shows $/resource on the Y-axis and resource usage on the X-axis (this can be for a single resource, or for multiple resources), a user defined upper bound Window, indicating that the user under no circumstances wishes to increase their resource usage below this Window, and a Minimum Savings Line, indicating that $/resource must be below a certain value otherwise increasing usage will cost too much. The Window can be described as a percentage or as a number of units from the actual usage. The Window and the Minimum

Savings Line both act as filters. In scenario A the $/resource rises as resource usage is increased, and quickly rises above the Minimum Savings Line, and doesn't drop down again until well beyond the upper bound window. In Scenario B the $/resource initially rises above the Minimum Savings Line but then quickly drops below it. For Scenario A, the cost sensitivity graph indicates that resource usage shouldn't be increased as the $/resource increases. For Scenario B, the cost sensitivity graph indicates that resource usage may be increased, but must be increased beyond point 721 to gain a desired $/resource level.

[0074] Price elasticity can also be calculated mid logging interval, based on actual resource usage and using a predictive curve to represent the remainder of the interval. An automated system may then take some action to increase or decrease resource usage based on the output.

[0075] It will be appreciated that although the foregoing explanation of cost sensitivity analysis was described in a pictorial format, that such evaluations may be performed by a cost sensitivity software module which would automate the process of determining the best opportunities. In one example, the cost sensitivity module would calculate the slope based on the window as defined by the user or the application.

[0076] Once the cost curves have been evaluated, the third step is to sort the top opportunities. These are: opportunities to use less resource and save money, opportunities to use more resource and not pay a lot more for it, or pay less per resource, and opportunities to not change current usage because a change (either increase or decrease) will cost more in total or per resource. Sorting the opportunities can be done based on one or more criteria, such as angle of slope, distance of change from actual and magnitude of usage at 100%. The steepness of the slope indicates the cost savings / added expense related to a change. For example, a steep negative slope is associated with reduced cost, a steep positive slope is associated with increased cost, and a flat slope is associated with an unchanging cost. Distance from actual weighs the value of the change based on how realistic it is. For example, dropping 10% of usage may be realistic, whereas dropping 90% is likely less realistic. Magnitude of usage at 100% measures the net value of making a change, as when the magnitude of resource usage is low there is less value obtained from reducing usage as when magnitude of usage is high. The output

of the sorting may be a table listing intervals and change in resource usage required for each interval to result in a certain cost saving.

Other Uses

[0077] It will be appreciated that the disclosed embodiments are not limited to uses in energy systems, but may also be used to distribute charges associated with any utility where usage can be frequently measured, such as water, air, gas, steam, emissions, Bandwidth, and computer processing availability/power (typically measured in Million Instructions Per Second ("MIPS")).

[0078] Where the utility being distributed is emissions, utility resources can be things like energy usage and energy type (coal, hydro etc), and the rate data can be some tariff used to determine emissions charges. The rate data may include production peaks, emissions budgets (how many emissions can be produced in a given time frame) and TOU penalties, for example, TOU may be important for smog emissions. In the simplest model, the charges can be distributed based on energy used. The output may be used for various analysis, for example what-if the energy source changes, what will the effect on emissions be, or what-if the weather changes, what will the emissions be.

[0079] Where the utility being distributed is computer processing availability, such as in an "on-demand" computing environment, the utility resources can be CPU, disk space, database space, web services, bandwidth, MIPS, terabytes of storage, number of transactions, and so forth. The rate data could be some tariff used to determine computer processing charges. The rate data may include a TOU component, where the computing resources are more expensive to use when other users want to use them as well.

[0080] On-demand computing is an increasingly popular enterprise model in which computing resources are made available to the user as needed. The resources may be maintained within the user's enterprise, or made available by a service provider. The on-demand model was developed to overcome the common challenge to an enterprise of being able to meet fluctuating demands efficiently. Because an enterprise's demand on computing resources can vary drastically from one time to another, maintaining sufficient resources to meet peak requirements can be costly. Conversely, if the enterprise cuts costs by only maintaining minimal computing resources, there will not be sufficient

resources to meet peak requirements. The above disclosed embodiments may be used in an on-demand computing environment to monitor usage, consumption and/or costs as described above to achieve increased efficiencies and cost savings.

[0081] It is therefore intended that the foregoing detailed description be regarded as illustrative rather than limiting, and that it be understood that it is the following claims, including all equivalents, that are intended to define the spirit and scope of this invention.